

Attorney's Docket No.: 06666/078001/USC-3022

REMARKS

Reconsideration and allowance of the above referenced application are respectfully requested.

Claims 7 stands rejected under 35 USC 112, first paragraph, as allegedly not being supported. The rejection alleges that the specification does not describe how a discourse tree can be derived for any input text segment. In response, the term "segment" has been removed from the claim. As amended, any discourse tree that has its nodes labeled with discourse relations and rhetorical status e.g., nucleus or satellites, can be reconstructed. This can be done starting from a sequence of elementary discourse units via the incremental application of a sequence of reduce and shift operations.

Claims 1-6 and 8-29 stand rejected under 35 USC 102 as allegedly being anticipated by Dr. Marcu's Ph.D. thesis. However, this contention has been obviated by the amendment of the claims to emphasize that the rules are automatically created by the machine. In contrast, the thesis work describes rules that were manually created by Dr. Marcu, based on introspection and reading of the training set.

The thesis work describes rules created manually, by Dr. Marcu, using introspection and data analysis of cue-phrase usages. In contrast, the claims, e.g., claim 1, describe rules

Attorney's Docket No.: 06666/078001/USC-3022

that are learned automatically from a set of discourse trees. Claim 1, for example, defines "using the computer to automatically learn a set of... rules." This was not done in Dr. Marcu's thesis. The training corpus, in fact, was created especially for carrying out the work described in the patent; it did not even exist when the Ph.D. thesis was written.

More specifically, the thesis shows how starting from a corpus of texts that contain many instances of cue phrases (phrases such as "because", "in order to", etc) -- see Section 4.4 in the Ph.D. thesis --, can:

1. manually annotate and store in a database the discourse usages of the cue phrases (Sections 4.4.2-3, pages 119-130) in order to develop an understanding of the types of discourse relations signaled by them; and

2. manually determine, via a process of introspection and analysis of the annotated data, the programming actions that a discourse parser may have to take when such cue phrases are found in text (Section 4.4.4, pages 130-132).

All of these steps are manual, as mentioned explicitly in Section 4.4.4, page 130: "Once the database had been created, I analyzed each record in it and updated its fields according to the requirements described in section 4.4.3. Tables 4.2 and 4.3 show the information that I associated with the fields when I analyzed the text fragments shown in (4.13) and (4.14)

Attorney's Docket No.: 06666/078001/USC-3022

respectively."

Chapter 5, page 145 also makes this explicit: "I used the cue phrases and the orthographic environments that characterized the cue phrases of the first two classes in order to manually develop a set of regular expressions that can be used to recognize potential discourse markers in naturally occurring texts."

The same is true about the component that hypothesizes rhetorical relations (chapter 5, page 160): "In order to hypothesize rhetorical relations, I manually associated with each of the regular expressions that can be used to recognize potential discourse markers in naturally occurring texts (see section 5.2.1) a set of features for each of the discourse functions that a regular expression can signal."

Therefore, all these steps were manual. There was no disclosure of automatic learning, as defined by claim 1.

Claim 4 defines iterative operations on the discourse units. This claim is based on a fundamental difference between the claim 4 system, and the thesis.

The algorithms for deriving a discourse structure described in the Ph.D. thesis are traditional bottom-up parsing algorithms (chart-parsing and CKY style). In contrast, the algorithm described in the disclosure is incremental. The embodiment describes left-to-right operations.

Attorney's Docket No.: 06666/078001/USC-3022

More specifically, the chart parsing and CKY algorithms in the thesis construct a discourse parse tree in a bottom-up fashion, recursively for increasingly larger textual units (sentences, paragraphs, sections, texts) - see Figure 5.1, Ph.D. thesis, page 142 and the high-level description of the algorithm in Section 5.1.2 (pages 142-143). The section of the thesis that describes in detail the parsing process is that between pages 177-185.

The discourse parsing process of claim 4 is completely different as it allows iteratively performing the operations, to incrementally build the segments. This would not have been possible using the thesis teachings, which required increasingly larger textual units (sentences, paragraphs, sections, texts), but it rather goes from the beginning-of-a-text to the end-of-a-text by keeping track of decisions made at all levels (sentences, paragraphs, sections, texts) concurrently.

Claim 7 defines a discourse tree, and Claims 8-11 define additional aspects thereof. Nothing in the thesis suggests such trees.

Claim 10 defines a stack that keeps track of partial, incomplete discourse parse trees. Again, this is not taught or suggested by the thesis.

Claim 16 has been amended to emphasize the automatic generation by the computer, and should be allowable for

Attorney's Docket No.: 06666/078001/USC-3022

analogous reasons to those discussed above, along with the claims that depend therefrom.


Claim 21 has also been analogously amended, and should be allowable for the reasons discussed above. Claim 21 also defines the combination of the elementary discourse units into larger discourse tree units, the patentability of which has been described above. Claim 27 has been similarly amended, and defines not only the automatic building, but also the stack which holds the discourse tree segments. As described above, this was not suggested by the thesis.

It is believed that all of the pending claims have been addressed in this paper. However, failure to address a specific rejection, issue or comment, does not signify agreement with or concession of that rejection, issue or comment. In addition, because the arguments made above are not intended to be exhaustive, there may be reasons for patentability of any or all pending claims (or other claims) that have not been expressed. Finally, nothing in this paper should be construed as an intent to concede any issue with regard to any claim, except as specifically stated in this paper, and the amendment of any claim does not necessarily signify concession of unpatentability of the claim prior to its amendment.

Attorney's Docket No.: 06666/078001/USC-3022

Applicant asks that all claims be allowed. No fee is believed to be due. However, please apply any applicable charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: December 9, 2005

Scott C. Harris
Reg. No. 32,030

Fish & Richardson P.C.
PTO Customer No. 20985
12390 El Camino Real
San Diego, California 92130
(858) 678-5070 telephone
(858) 678-5099 facsimile

10577116.doc